Ling/CSE 472: Introduction to Computational Linguistics

5/23/17

Compositional Semantics

Overview

- The Conduit Metaphor and speaker meaning v. sentence meaning
- Compositionality and the role of syntax
- Compositionality in the Grammar Matrix
- ‘Layers of Interpretation: On Grammar and Compositionality’
- Reading questions

The conduit metaphor (Reddy 1979)

- “Your thoughts here don’t quite make it across.”
- “It is very difficult to put this concept into words.”
- “Never load a sentence with more thought than it can carry.”
- “The passage conveys a feeling of excitement.”
- “John says that he cannot find your idea anywhere in the passage.”
- “Get your insights down on paper at once.”
- “The concept made its way very quickly into the universities.”

Conduit metaphor, “major framework” (Reddy 1979: 290)

1. language functions like a conduit, transferring thoughts bodily from one person to another

2. in writing and speaking, people insert their thoughts or feelings in the words

3. words accomplish the transfer by containing the thoughts or feelings and conveying them to others

4. in listening or reading, people extract the thoughts and feelings once again from the words

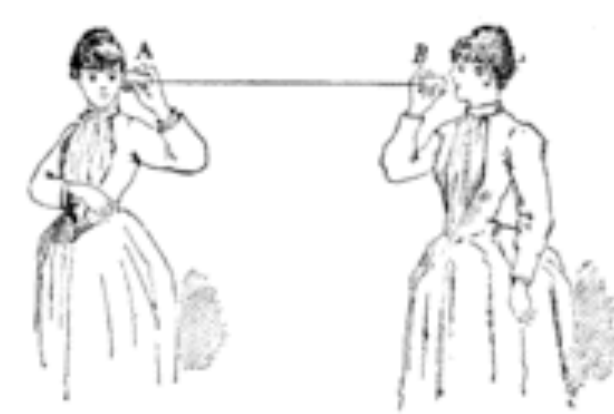


Fig. 76. Trådtelefon.

Conduit metaphor, “minor framework” (Reddy 1979: 291)

1. thoughts and feelings are ejected by simply speaking them into an external "idea space"

2. thoughts and feelings are reified in this external space, so that they exist independently of any need for living human beings to think or feel them

3. these reified thoughts and feelings may, or may not, find their way back into the heads of living humans

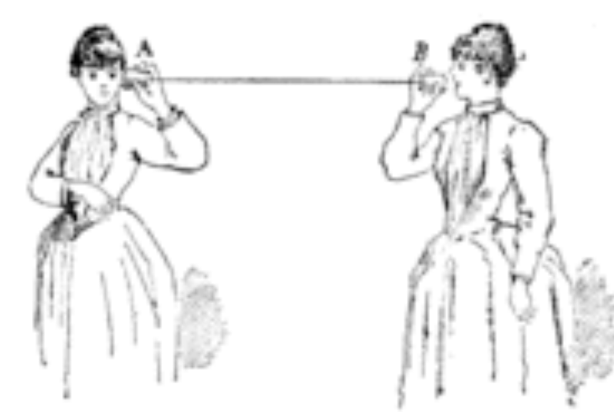


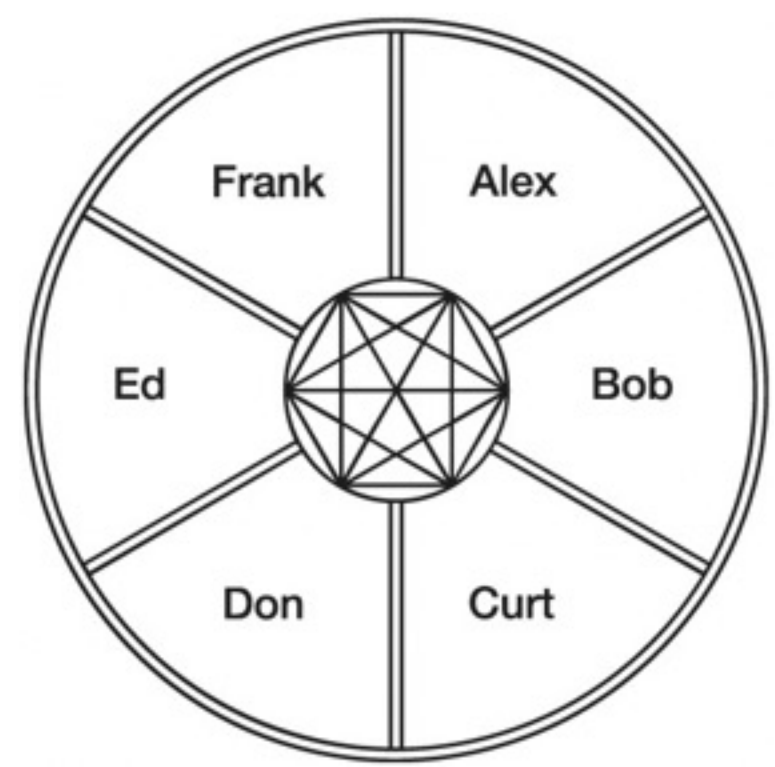
FIG. 76. Trådtelefon.

Toolmaker's paradigm/radical subjectivism (Reddy 1979)

1. Each person has their own environment, with their own "indigenous materials"

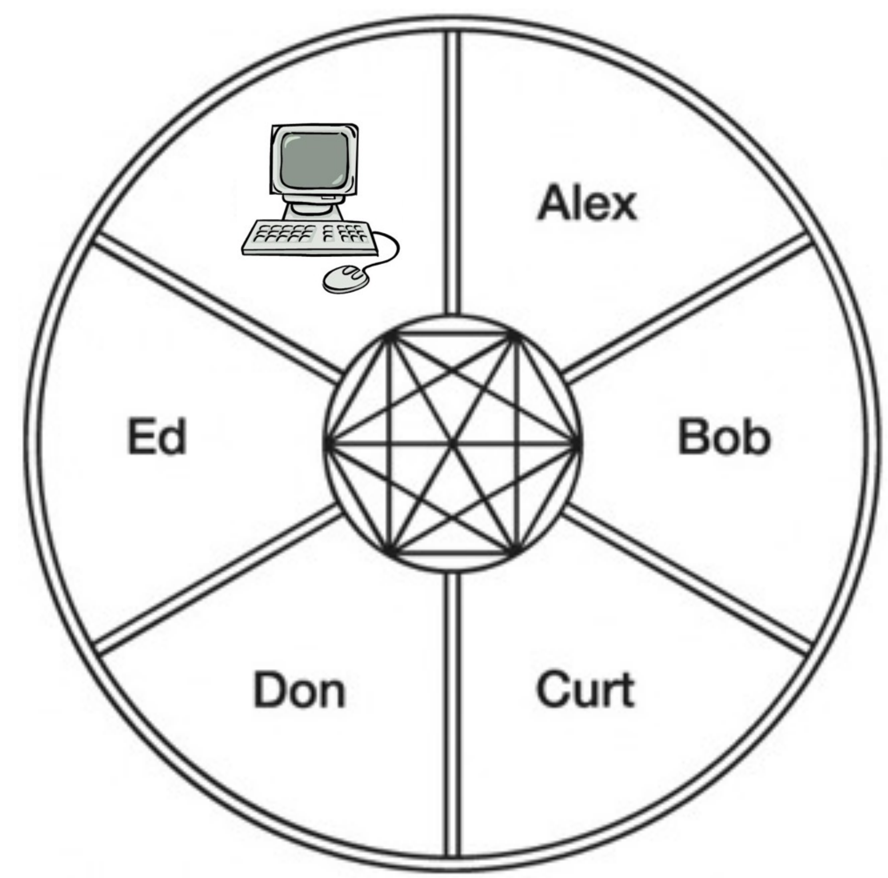
2. Communication involves using language (and other signals) as clues to construct private representations from indigenous materials

3. Each interlocutor's representations will be scattered (different) unless they expend energy to coordinate them



Questions to consider

- In what ways do NLP/NLU technologies incorporate the conduit metaphor, and how might this affect their success?
- Given that machines have very different tools in their own “indigenous environments” how can they effectively join communication as understood in the toolmakers paradigm?



Clark 1996: *Using Language*

- Language use is a *joint action*: “Alan intends Barbara to recognize that he wants her to say whether or not she saw his dog run by on the sidewalk, and she is to see this in part by recognizing that intention. The remarkable thing about Alan’s intentions is that they involve Barbara’s thoughts about those very intentions.” (p.12)

- “Words and sentences are *types* of signals, linguistic units abstracted away from any occasion on which they might be used, stripped of all relation to particular speakers, listeners, times, and places.

- “To describe them is to describe the conventions for their use within speech communities.

- “But utterances are the actions of producing words, sentences, and other things on particular occasions by particular speakers for particular purposes.” (p.128)

Reconciling Clark and Reddy

- Through experience within our speech communities, we learn (and help create) shared linguistic conventions.
- These conventions support fairly consistent calculation of ***sentence meaning*** by different speakers in the same community.
- The sentence meaning of an utterance (together with its form) serves as a clue which a toolmaker-listener can use to construct his/her representation of the toolmaker-speaker's ***speaker meaning***

(See also Grice 1968)

Overview

- The Conduit Metaphor and speaker meaning v. sentence meaning
- Compositionality and the role of syntax
- Compositionality in the Grammar Matrix
- ‘Layers of Interpretation: On Grammar and Compositionality’
- Reading questions

Recovering sentence & speaker meaning: Semantic dependencies

- Semantic dependencies form a central part of sentence meaning
- Semantic dependencies are key to getting the most from sentence meaning when aiming to construct a representation of speaker meaning

先生によると男の子よりも女の子がポケモンがすきだ。[jpn]

先生 によると 男の子 よりも 女の子 が ポケモン が すき だ。

teacher (.) boy (.) girl (.) Pokemon (.) like (.)

teacher ACCORDING.TO boy THAN girl NOM Pokemon NOM like COP.PRES

according-to(e4, e3, x6)
teachers(x6)
like(e3, x14, x17)

girls(x14)
pokemon(x17)
more-than(e22, x14, x23)
boys(x23)

Overview

- The Conduit Metaphor and speaker meaning v. sentence meaning
- Compositionality and the role of syntax
- Compositionality in the Grammar Matrix
- ‘Layers of Interpretation: On Grammar and Compositionality’
- Reading questions

Example: Grammar Matrix semantic compositionality principle

```
basic-binary-phrase := phrase &
  [ SYNSEM.LOCAL.CONT [ RELS [ LIST #first,
                         LAST #last ],
             HCONS [ LIST #scfirst,
                 LAST #sclast ],
             ICONS [ LIST #icfirst,
                 LAST #iclast ] ],
    C-CONT [ RELS [ LIST #middle2,
             LAST #last ],
           HCONS [ LIST #scmiddle2,
               LAST #sclast ],
           ICONS [ LIST #icmiddle2,
               LAST #iclast ] ],
    ARGS < sign & [ SYNSEM.LOCAL local &
                  [ CONT [ RELS [ LIST #first,
                               LAST #middle1 ],
                           HCONS [ LIST #scfirst,
                               LAST #scmiddle1 ],
                           ICONS [ LIST #icfirst,
                               LAST #icmiddle1 ] ] ] ],
        sign & [ SYNSEM.LOCAL local &
                  [ CONT [ RELS [ LIST #middle1,
                               LAST #middle2 ],
                           HCONS [ LIST #scmiddle1,
                               LAST #scmiddle2 ],
                           ICONS [ LIST #icmiddle1,
                               LAST #icmiddle2 ] ] ] ] > ].
```

Example: Grammar Matrix head-complement structures

```
basic-head-1st-comp-phrase := basic-head-comp-phrase &
  [ SYNSEM.LOCAL.CAT.VAL.COMPS #comps,
    HEAD-DTR.SYNSEM.LOCAL.CAT.VAL.COMPS < #synsem . #comps >,
    NON-HEAD-DTR.SYNSEM #synsem ].
```

```
transitive-lex-item := basic-two-arg-no-hcons & basic-icons-lex-item &
  [ ARG-ST < [ LOCAL.CONT.HOOK [ INDEX ref-ind & #ind1,
                                 ICONS-KEY.IARG1 #clause ] ],
             [ LOCAL.CONT.HOOK [ INDEX ref-ind & #ind2,
                                 ICONS-KEY.IARG1 #clause ] ] >,
    SYNSEM [ LKEYS.KEYREL [ ARG1 #ind1,
                            ARG2 #ind2 ],
             LOCAL.CONT.HOOK.CLAUSE-KEY #clause ] ].
```

Overview

- The Conduit Metaphor and speaker meaning v. sentence meaning
- Compositionality and the role of syntax
- Compositionality in the Grammar Matrix
- ‘Layers of Interpretation: On Grammar and Compositionality’
- Reading questions

Layers of Interpretation: On Grammar and Compositionality

Emily M. Bender, Dan Flickinger, Stephan Oepen,
Woodley Packard & Ann Copestake

IWCS 2015
17 April 2015
Queen Mary College of London

Central claims

- Sentence meaning, but not speaker meaning, is compositional
- Systems attempting to understand speaker meaning would benefit from reusable, automatically derivable, task-independent representations of sentence meaning
- A compositional approach to creating sentence meaning representations provides
  - Comprehensiveness
  - Consistency
  - Scalability

Overview

- Introduction: Sentence meaning representation in NLU
- Compositionality: Working definition, survey of semantic annotations
- One approach: The English Resource Grammar (ERG)
- Benefits of Compositionality
- Inter-Annotator Agreement
- How to get the best of both worlds

‘Semantic parsing’

- “the task of mapping natural language sentences to complete formal representations which a computer can execute for some domain-specific application” (Kate and Wong, 2010)
- Not really new: Woods et al 1972, Gawron et al 1982, Alshawi 1992, ...
- But much current work in semantic parsing is different:
  - Emphasis on task- or domain-specific representations
  - Expectation to map straight from surface strings, without explicitly modeling syntax

Sentence meaning, speaker meaning

- Learning correlations between domain-typical surface forms and task-specific representations conflates:
  - timeless/conventional/standing/sentence meaning
  - utterer/occasion/speaker meaning
- Drawbacks:
  - resolving the same problems around grammatical structure for each task
  - unlikely to scale to general-purpose NLU

Leveraging sentence meaning

- Machines don't have access to any direct representation of speaker meaning, only to natural language utterances
- Sentence meaning doesn't determine situated speaker meaning, but is an important cue to it (Quine 1960, Grice 1968, Ready 1979, Clark 1996)
- A task-independent, comprehensive representation of sentence meaning capturing exactly the information in the linguistic signal itself should benefit NLU systems

Overview

- Introduction: Sentence meaning representation in NLU
- Compositionality: Working definition, survey of semantic annotations
- One approach: The English Resource Grammar (ERG)
- Benefits of Compositionality
- Inter-Annotator Agreement
- How to get the best of both worlds

A meaning representation system is compositional if (working definition):

- it is grounded in a finite (possibly large) number of atomic symbol-meaning pairings
- it is possible to create larger symbol-meaning pairings by combining the atomic pairings through a finite set of rules;
- the meaning of any non-atomic symbol-meaning pairing is a function of its parts and the way they are combined;
- this function is possibly complex, containing special cases for special types of syntactic combination, but only draws on the immediate constituents and any semantic contribution of the rule combining them; and
- further processing will not need to destructively change a meaning representation created in this way to create another of the same type.

Semantic annotation survey: Compositional layer

- Predicate-argument structure
- Partial constraints on:
  - Scope of negation and other operators
  - Restriction of quantifiers
  - Modality
  - Tense/aspect/mood
  - Information structure
  - Discourse status of referents of NPs
  - Politeness
- Possibly compositional, but not according to sentence grammar:
  - Coherence relations/ rhetorical structure

Semantic annotation survey: Sub-compositional layer

- Fine-grained word-sense annotations; entity types; links to ontologies
  - Concern only atoms
  - What is built compositionally is not the choice of atom, but the connections between them
  - Sense underspecification: Only discriminate between supersenses correlating with morphosyntactic differences

Semantic annotation survey: Computation on top of compositional backbone

- Quantifier scope ambiguity resolution (e.g. Higgins & Sadock 2003)
- Coreference resolution (e.g. Hobbs 1979)
- Determination of the focus of negation (e.g. Blanco and Moldovan 2011)
    - Build on partial constraints provided by the grammar
    - Not compositional in that it's never (strictly) constrained by grammatical structure

Semantic annotation survey: Discourse processing

- Presupposition projection (e.g. van der Sandt 1992, Zaenen & Karttunen 2013, Venhuizen et al 2013)

- Coherence relations/rhetorical structure (e.g. Marcu 1997)

- Discourse moves/adjacency pairs (e.g. Shriberg et al 2004)

  - Again, build on information provided during sentence-level processing

  - Concerns cross-sentential relations, so not compositional according to sentence grammar

  - Open question: compositional processes at higher levels of structure?

Semantic annotation survey: Doing things with words

- Hedge detection (e.g. Vincze et al 2008)
- Authority claim, alignment move, and other social act annotation (e.g. Morgan et al 2013)
- Pursuit of power in dialogue (e.g. Swayamdipta & Rambow 2012)
  - Not anchored in the structure of sentences
  - Relate to speaker goals

Overview

- Introduction: Sentence meaning representation in NLU
- Compositionality: Working definition, survey of semantic annotations
- One approach: The English Resource Grammar (ERG)
- Benefits of Compositionality
- Inter-Annotator Agreement
- How to get the best of both worlds

ERG: The English Resource Grammar (Flickinger 2000, 2011)

- Under continuous development since 1993
- Framework: Head-driven Phrase Structure Grammar (Pollard & Sag 1994)
- 1214 release: 225 syntactic rules, 70 lexical rules, 975 leaf lexical types
- Open-source and compatible with open-source DELPH-IN processing engines (www.delph-in.net)
- Broad-coverage: 85-95% on varied domains: newspaper text, Wikipedia, bio-medial research literature (Flickinger et al 2010, 2012; Adolphs et al 2008)
  - ‘Bridging’ rule analyses enable 100% coverage
- Output: derivation trees paired with meaning representations in the Minimal Recursion Semantics framework---English Resource Semantics (ERS)

Redwoods: ERG-based treebanking (sembanking) (Oepen et al 2004)

- Minimal discriminants (Carter 1997): Properties of derivation trees partitioning parse forest per item
- Allows annotators to swiftly navigate even very large parse forests to select intended analysis or reject all analyses
  - 37,200 words of the Brown corpus annotated in 1400 minutes (1.7 sentences/min)
- All annotation decisions are recorded and can be rerun against updated parse forests produced by updated grammar versions
- Current Redwoods release (7th growth) includes 45,000 sentences of annotated text across genres including Wikipedia, tourism brochures, ...
- Analyses can be viewed as full HPSG analyses, ERS only, or even simpler syntactic or semantic dependency representations

Overview

- Introduction: Sentence meaning representation in NLU
- Compositionality: Working definition, survey of semantic annotations
- One approach: The English Resource Grammar (ERG)
- Benefits of Compositionality
- Inter-Annotator Agreement
- How to get the best of both worlds

Why a grammar-based compositional approach?

- Argued above for the importance of task-independent, sentence-meaning annotations
- Can created be done:
  - Non-compositionally, as in Abstract Meaning Representation (AMR; Langkilde & Knight 1998, Banarescu et al 2013)
  - Compositionally, by hand, as in PropBank (Kingsbury & Palmer 2002) and FrameNet (Baker et al 1998)
  - Compositionally, with a machine-readable grammar, as in Redwoods (Oepen et al 2004), TREPIL (Rosén et al 2005), or the Groningen Meaning Bank (Basile et al 2012)

Benefits of compositionality: Comprehensiveness

- Grammar-based compositional approach ⇒ Every word and syntactic structure must be accounted for, or specifically deemed semantically void
- Narrower paraphrase sets, compare AMR (1), (2) (Banarescu et al 2014) to ERS (3)

(1)
a. No one ate.
b. Every person failed to eat.

(2)
a. The boy is responsible for the work.
b. The boy is responsible for doing the work.
c. The boy has the responsibility for the work.

Benefits of compositionality: Comprehensiveness

- Grammar-based compositional approach ⇒ Every word and syntactic structure must be accounted for, or specifically deemed semantically void

- Narrower paraphrase sets, compare AMR (1), (2) (Banarescu et al 2014) to ERS (3)

(3)
  a. Kim thinks Sandy gave the book to Pat.
  b. Kim thinks that Sandy gave the book to Pat.
  c. Kim thinks Sandy gave Pat the book.
  d. Kim thinks the book was given to Pat by Sandy.
  e. The book, Kim thinks Sandy gave to Pat.

Benefits of compositionality: Comprehensiveness

- Task-independent semantic representations can't abstract away from seemingly less relevant nuances of sentence meaning

- Compositional approach facilitates capturing more detail

$$\langle\, h_1,\; \left|\begin{array}{l} h_4\text{:person}\langle 0\!:\!6\rangle(\text{ARG0 } x_5), \\ h_6\text{:\_no\_q}\langle 0\!:\!6\rangle(\text{ARG0 } x_5, \text{ RSTR } h_7, \text{ BODY } h_8), \\ h_2\text{:\_eat\_v\_1}\langle 7\!:\!11\rangle(\text{ARG0 } e_3, \text{ ARG1 } x_5, \text{ ARG2 } i_9) \end{array}\right| \; \{\, h_1 =_q h_2,\ h_7 =_q h_4 \,\}\,\rangle$$

```
(e / eat-01
 :polarity -
 :ARG0 (p / person
   :mod (e / every)))
```

$$\langle\, h_1,\; \left|\begin{array}{l} h_4\text{:\_every\_q}\langle 0\!:\!5\rangle(\text{ARG0 } x_6, \text{ RSTR } h_7, \text{ BODY } h_5), \\ h_8\text{:\_person\_n\_1}\langle 6\!:\!12\rangle(\text{ARG0 } x_6), \\ h_2\text{:\_fail\_v\_1}\langle 13\!:\!19\rangle(\text{ARG0 } e_3, \text{ ARG1 } h_9), \\ h_{10}\text{:\_eat\_v\_1}\langle 23\!:\!27\rangle(\text{ARG0 } e_{11}, \text{ ARG1 } x_6, \text{ ARG2 } i_{12}) \end{array}\right| \; \{\, h_1 =_q h_2,\ h_7 =_q h_8,\ h_9 =_q h_{10} \,\}\,\rangle$$

Benefits of Compositionality: Consistency

- Requiring meaning representations to be grounded in both the lexical items and syntactic structure of the strings being annotated significantly reduces the space of possible annotations

- Grammar based approach allows encoding of design decisions for machine application

  - Ex: arguments of *when*

- Human input still required, but choosing among representations is far simpler than authoring them

  - Development of grammar is still a big investment, but with big returns as the same grammar is applied over more and more text

Benefits of Compositionality: Scalability

- In amount of text annotated: Initial development of grammar pays off as it is applied to as much text as desired
- In genre diversity of the resource: One and the same grammar can be applied to texts from multiple different domains
  - Robustness techniques can compensate for lack of grammar coverage
- In the complexity of the annotations themselves: Grammar updates can be efficiently propagated across the treebank by reparsing corpus and rerunning annotation decisions (Oepen et al 2004)
  - Improve analyses of particular phenomena, or add layers of grammar-based annotation (e.g. partial constraints on information structure)

Overview

- Introduction: Sentence meaning representation in NLU
- Compositionality: Working definition, survey of semantic annotations
- One approach: The English Resource Grammar (ERG)
- Benefits of Compositionality
- Inter-Annotator Agreement
- How to get the best of both worlds

Inter-Annotator Agreement study

- Data source: Sentences sampled from Antoine de Saint Exupéry's *The Little Prince*
- Three expert annotators
- Annotated 50-sentence trial set, then adjudicated, updating annotation guidelines as indicated
- Annotated 150-sentence sample set, then measured IAA, then produced adjudicated gold standard
- Repeat above steps with 'bridging' analyses in

Agreement Metrics

- NB: Chance-corrected IAA measures as yet unavailable for graph-structured annotations
- Exact match: Full ERS identical between annotators
- Elementary Dependency Match (Dridan & Oepen 2011)
  - Computed over sets of triples from reduction of ERS to Elementary Dependency Structures (EDS)
  - EDMa: Argument identification only
  - EDMna: Argument identification + predicate name identification

IAA Results

	Annotator Comparison			
Metric	**A vs. B**	**A vs. C**	**B vs. C**	**Average**
Exact Match	0.73	0.65	0.70	0.70
EDM_a	0.93	0.92	0.94	0.93
EDM_{na}	0.94	0.94	0.95	0.94

- Compare Banarescu et al (2013) triple-based IAA for AMR over web text of 0.71

Overview

- Introduction: Sentence meaning representation in NLU
- Compositionality: Working definition, survey of semantic annotations
- One approach: The English Resource Grammar (ERG)
- Benefits of Compositionality
- Inter-Annotator Agreement
- How to get the best of both worlds

The best of both worlds

- Grammar-based compositionally created meaning representations can only access information that can be implemented in a grammar

- But having comprehensive, integrated meaning representations is desirable (Basile et al 2012, Banarescu et al 2013), Ide & Suderman 2007)

- Proposal: Start from grammar-based meaning representations and process to produce larger paraphrase sets and/or add layers, such as fine-grained word sense annotations, additional coreference links, cross-sentential discourse relations

Central claims (reprise)

- Sentence meaning, but not speaker meaning, is compositional
- Systems attempting to understand speaker meaning would benefit from reusable, automatically derivable, task-independent representations of sentence meaning
- A compositional approach to creating sentence meaning representations provides
  - Comprehensiveness
  - Consistency
  - Scalability

Thank you

Reading questions

- It's my understanding that dependence on grammar rules and rule-to-rule lexical entries will rule out ungrammatical inputs. However, humans have intuition/pragmatism about the meaning and intention of certain ungrammatical sentences (such as those uttered by young children). Would new sets of rules created to stay consistent with the rule-to-rule hypothesis, or are there statistical / non-grammar rule approaches to compositional semantics?

- Are there rule-to-rule based systems that work analogously to a probabilistic parser, where the probability for different semantic interpretations are used to pick the best candidate?

Reading questions

- In figure 18.4, Verb -> opened and Verb -> closed have very different semantic attachments even though they can be used in similar environments. I think that the semantic attachment for opened is meant only for the transitive case, and the semantic attachment for closed is meant only for the intransitive case. For a complete system, I would expect both the transitive and intransitive cases to be included, so does the system have to use the syntactic structure of the sentence to decide what semantic attachments to use?

Reading questions

- As I understand, in "holes semantics", holes are the places that other expressions can be "plugged in" under certain constraints. But what exactly are these holes and why are we using this approach?

- How exactly are the dominance constraints for the hope semantics approach chosen? The example in the book ("Every restaurant has a menu.") seems fairly simple, with only three "levels" of dominance; I imagine sentences can get more complex than that, but the book doesn't seem to make clear how those are handled.

Reading questions

- Is rule-to-rule method useful across languages?
- Semantic analyzer can make use of static knowledge from lexicon and grammar. Are there any kind of dynamic knowledge from lexicon and grammar?
- In the Early parser with semantic integration, what are some examples of semantic ill-formedness in a meaning representation that would cause its corresponding state to be blocked from entering the chart?

Reading questions

- What exactly is the scope of a quantifier? How does it affect the meaning of the sentence?
- How is sentences with EPP like "It is nice to meet you" going to be represented using the lambda-expression?
- Page 599 shows how we can express the statements of first-order logic and lambda calculus that we're using to capture semantic properties in terms of feature structures. Looking at it in the other direction, could we have used first-order logic and lambda calculus to express feature structures and unification?